Constructing the ASCI Grid

2000 NASA IPG Workshop
September 19-20, 2000

Steven Humphreys

Sandia National Laboratories

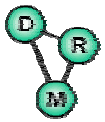
ASCI, The Accelerated Strategic Computing Initiative Program, includes participants from Los Alamos National Laboratory, Lawrence Livermore National Laboratory, and Sandia National Laboratories. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94AL85000.

Acknowledgments

- Esther Baldonado, SNL
- David Beck, SNL
- Judy Beiriger, SNL
- Hugh Bivens, SNL
- Rich Detry, SNL
- David Evensky, SNL
- Lee Ann Fisk, SNL
- Kathie Hiebert-Dodd, SNL
- Vic Holmes, SNL
- Wilbur Johnson, SNL
- Leslea Lehouq, SNL
- Kevin Long, SNL
- David Miller, SNL
- Donald McLaughlin, SNL
- Pat Moore, SNL
- Ronald Rhea, SNL
- Ruthe Vandewart, SNL
- Morris Jette, LLNL
- Keith Fitzgerald, LLNL
- Randal Rheinheimer, LANL
- the Globus project
- Grid Forum

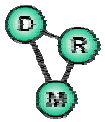
Outline

- Goals
- Approach
- Higher-Level Middleware
- Extensions to Globus
- Conclusions



Accelerated Strategic Computing Initiative (ASCI) Program Overview

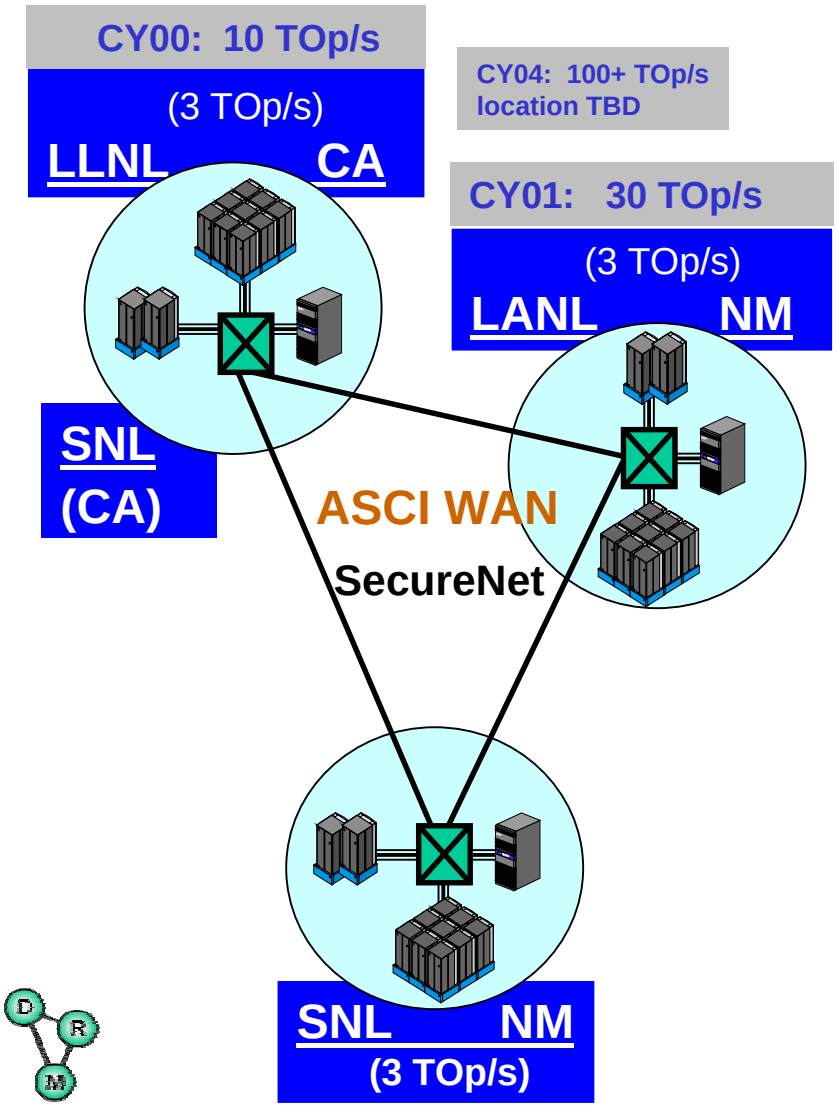
- An initiative of Defense Programs at the U.S. Department of Energy in collaboration with Lawrence Livermore National Laboratory, Los Alamos National Laboratory, and Sandia National Laboratories.
- The US commitment to ending underground nuclear testing calls for new means of verifying the safety, reliability, and performance of the aging US nuclear stockpile.
- By year 2010, applications must achieve high-resolution, three-dimensional, full-physics, and full-system capabilities.
- This level of simulation requires high-performance computing (HPC) far beyond our current level of performance.



ASCI Activities Within the Initiative

- Create seamless management: 1 program, 3 labs.
- Focus on advanced applications development.
  - Focus on 3-D, full-physics/systems applications.
  - Accelerate code performance.
  - Validate simulations via experiments and archival data.
- Focus on the high end of computing.
  - Accelerate the development of scalable architectures.
  - Develop partnerships with multiple vendors.
- Enable problem-solving environments.
  - Provide applications development support.
  - Ensure high-performance computing (HPC) access.
  - Develop tri-lab distributed computing environments.
- Encourage openness and collaboration.

Distance & Distributed Computing (DisCom²) is the Bridge Between Users and Computational Resources



Goal is Users Focus on Science:

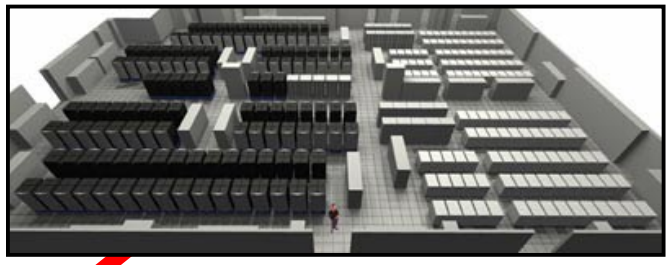
- Common access methods to resources.
- Hiding complexity of distributed system.
- Policy-based allocation.
- Routine sharing & aggregation of resources.
- Management of visualization, network, storage, data, and software resources.
- Resource- & application-independence.
- Domain-specific user environments.
- Collaborative tools.
- Coordinated use of collections of resources.

How:

- Data and simulation Integration.
- Distributed resource management.
- Networks.
- Security.

The ASCI Grid Will Start With A Production Distance Computing Environment FY01 Q2

ASCI Red

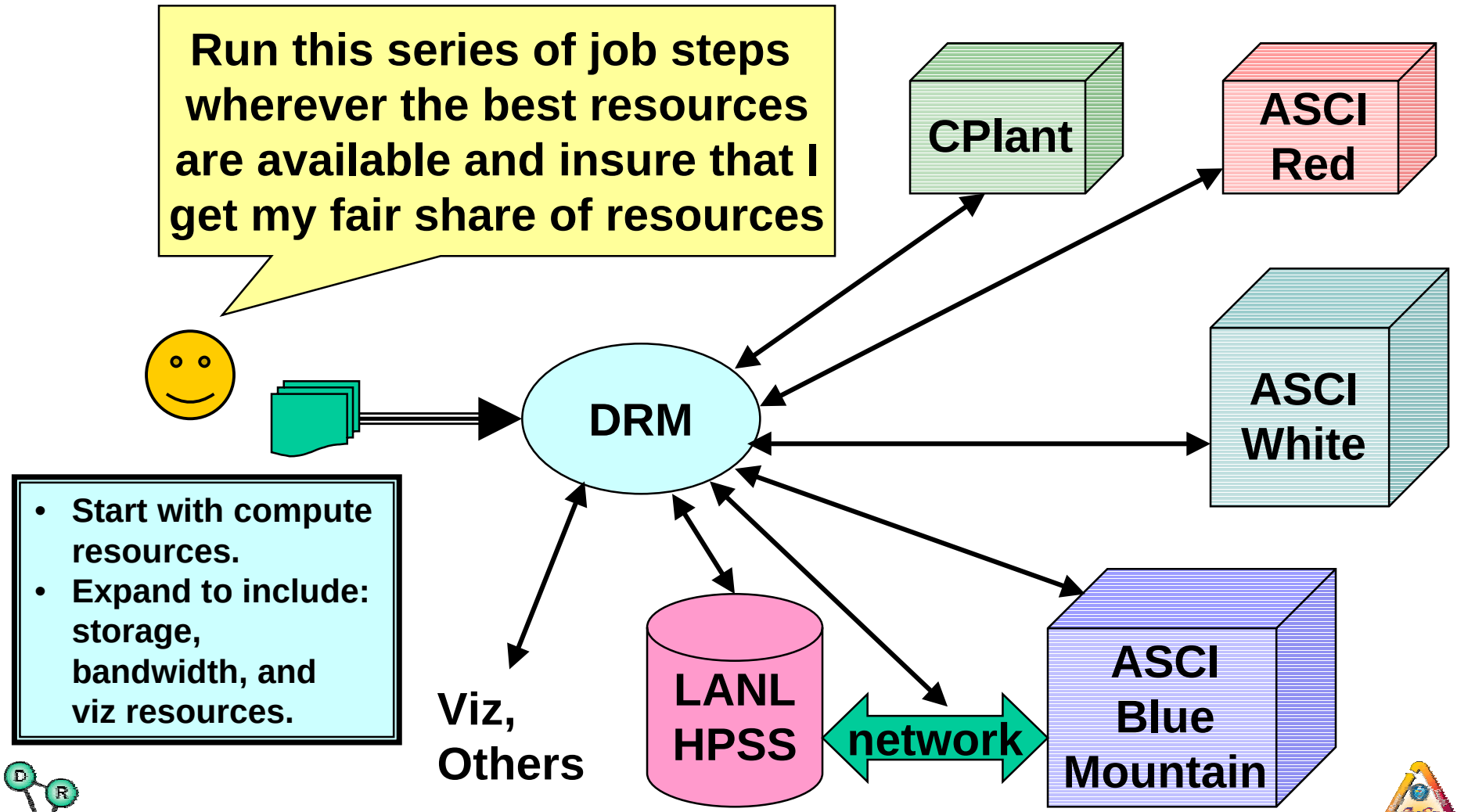


ASCI White

Distributed Resource Management Services on Top of Kerberos-Secured Globus Services

ASCI Blue Mountain

ASCI Will Deploy Distributed Services That Integrate All Resources Into an Effective Tri-Lab Grid FY03

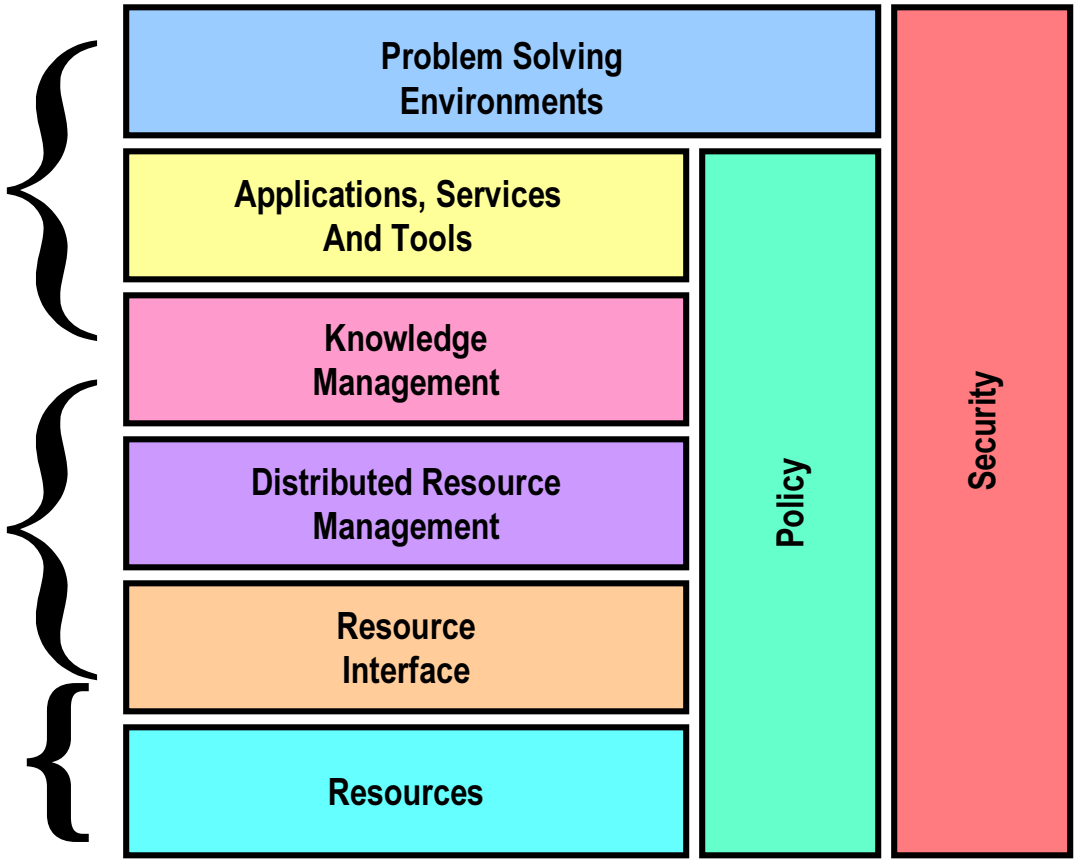


ASCI Grid Architecture Model

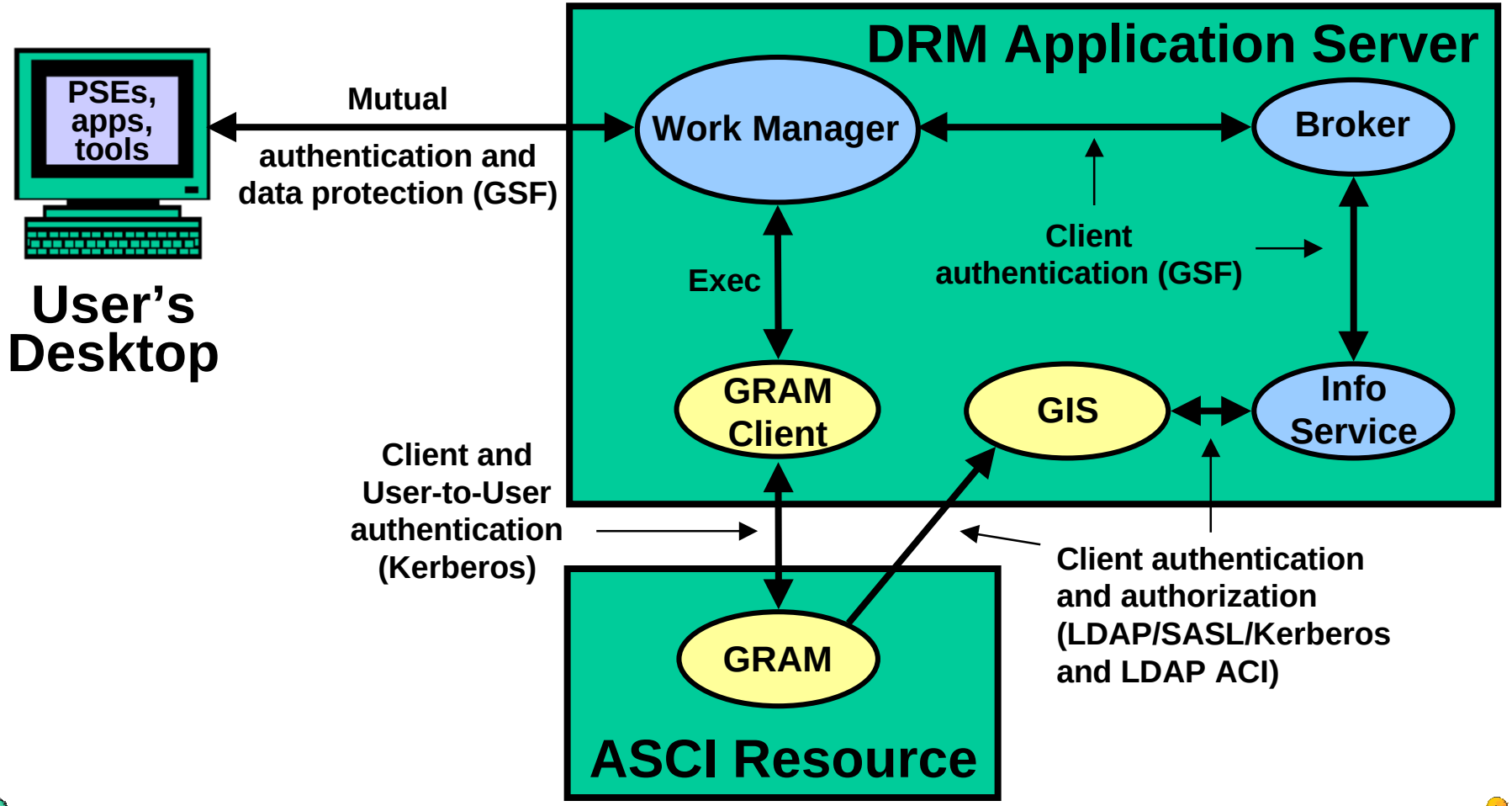
Problem Solving Environments allow users to focus on scientific task.

Grid Services provide discovery, reservation, allocation, monitoring, and control of resource collections.

Distributed resources are accessible regardless of location.



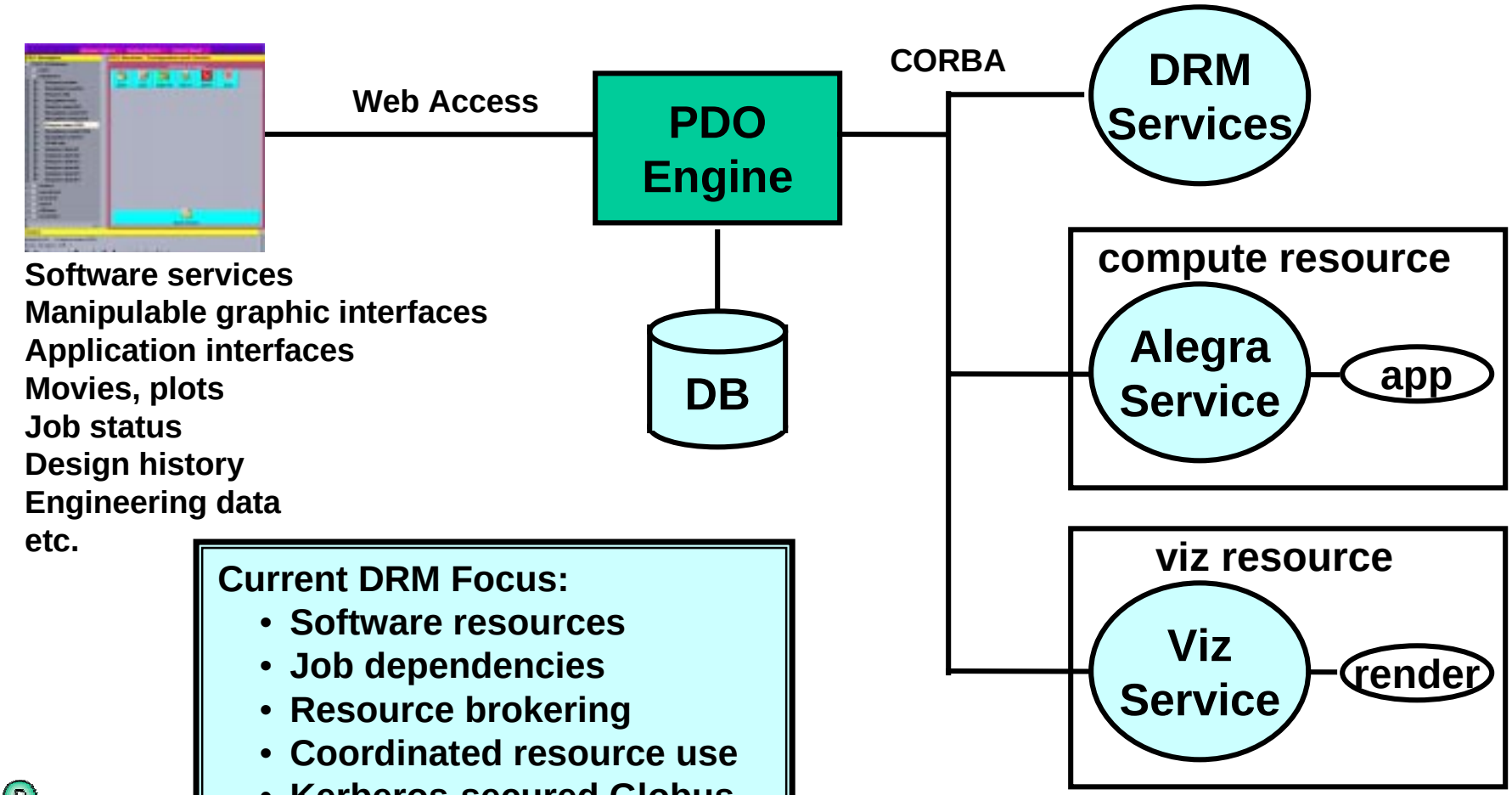
DRM Distributed Component Interactions



DRM Is Developing Middleware Services To Support Problem Solving Environments

Example: SI/PDO Product Design Environment

Software services
Manipulable graphic interfaces
Application interfaces
Movies, plots
Job status
Design history
Engineering data
etc.

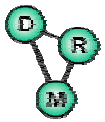


Current DRM Focus:

- Software resources
- Job dependencies
- Resource brokering
- Coordinated resource use
- Kerberos-secured Globus

Software Resources Provide Production Computing Services

- Computing Scenario
  - Codes used by different users at different sites.
  - Different versions must be available long-term.
  - Many resources are suitable for these jobs.
- Current Implementation
  - User can request “run code X anywhere.”
  - Software installed on resource maintained in Grid Info Service.
  - Availability of services monitored with Globus heartbeat monitor.
  - Application-specific user interfaces available in scripts, XML.



Next Steps for Software Resources

- **Software as a grid resource.**
- **Generalized approach for application-specific user interfaces.**
- **Support for distributing software throughout the grid.**

Work Management Services Support Job Dependencies and Task Sequencing

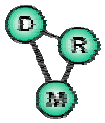
- **Current Implementation**
  - **CORBA *submit* method accesses Resource Broker and Globus services.**
  - **Common XML-based work request.**
  - **Coordinated job submission and file migration.**
  - **Sequential start-finish job dependencies.**
  - **dependencies.**
  - **Status and exception information.**

Next Steps for Work Management Services

- General grid event service.
- Concurrent access to multiple resources.
- Parallel and conditional dependencies.
- Better fault tolerance.

Brokering Services Support Resource Discovery and Criteria-Based Selection of Resources

- **Current Discovery Implementation**
  - **CORBA *acquireResources* method accesses Grid Info Services.**
  - **Attributes and constraints (e.g., specific machine, number of nodes, minimum memory, software).**
  - **Exclude specified resources.**
  - **Logical and conditional operations on criteria.**
  - **User authorization.**
- **Current Selection Implementation for “run Anywhere”**
  - **Select the host with lowest CPU load.**

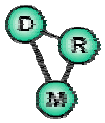


Next Steps for Brokering Services

- Additional “run anywhere” selection methods:
  - “Where job will finish soonest” based on jobs queued and machine capacity.
  - “Where my job will finish by <deadline>..”
- Support requests for collections of resources.
  - Compute,
  - Network,
  - Visualization,
  - ...

Globus MetaComputing Toolkit Forms The Basis Of DRM Services

- Globus provides foundation of grid services:
  - Capabilities match core ASCI requirements.
  - Side-by-side compatibility with present local resource managers.
- Changes needed for the ASCI environment:
  - Standalone Grid Information Service.
  - GRAM modifications for ASCI resources.
    - DPCS, CPlant™ PBS, NQS, LSF.
  - Monitoring service enhancements.
  - Kerberos V5 security via the GSSAPI.

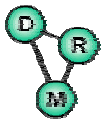


Kerberos Version 5 Security Services

- **Kerberos Version 5 is the authentication method used in the ASCI grid.**
  - **Kerberos-secured Globus services.**
  - **User-to-user authentication for MIT Kerberos library.**
  - **Access to LDAP server authenticated using a Netscape Directory Service (NDS) plug-in from PADL.**
  - **Access to LDAP server authorized using LDAP Access Control Instructions.**
  - **Sandia's Generalized Security Framework (GSF) used for authentication and data protection between DRM components.**
  - **Desktop access to grid services via grid-enabled applications (including generic job submission tools) or Secure Shell.**

Next Steps for Security Services

- **Security impact of deploying Globus 1.1.3 (or 1.1.4…):**
  - **New LDAP structure.**
  - **Pull vs. push model.**
- **Switch to OpenLDAP.**
  - **Present PADL solution considered temporary.**
  - **Waiting for Globus GIS authorization mechanism.**
- **Kerberos-authenticated web-enabled interface.**



Conclusions

- The Globus Grid Computing Toolkit has provided DRM with a solid foundation on which to build advanced services.
- Our present set of testbed services have already begun to demonstrate the utility of grid services in the ASCI environment.
- The DRM work management and resource brokering services enable new problem-solving methodologies that let users focus on science.
- With DOE accreditation of the DRM and Globus services, the ASCI production grid will be a world-class computational utility easily accessible by authorized users.